

Modularity bounds for clusters located by leading eigenvectors of the normalized modularity matrix

Dario Fasino¹

*Department of Chemistry, Physics, and Environment
University of Udine, Udine, Italy.*

Francesco Tudisco²

*Department of Mathematics and Computer Science,
Saarland University, Saarbrücken, Germany*

Abstract

Nodal theorems for generalized modularity matrices ensure that the cluster located by the positive entries of the leading eigenvector of various modularity matrices induces a connected subgraph. In this paper we obtain lower bounds for the modularity of that set of nodes showing that, under certain conditions, the nodal domains induced by eigenvectors corresponding to highly positive eigenvalues of the normalized modularity matrix have indeed positive modularity, that is they can be recognized as modules inside the network. Moreover we establish Cheeger-type inequalities for the cut-modularity of the graph, providing a theoretical support to the common understanding that highly positive eigenvalues of modularity matrices are related with the possibility of subdividing a network into communities.

Keywords: Nodal domain; community detection; modularity; Cheeger inequality

2010 MSC: 05C50, 15A18, 15B99

1. Introduction

The study of community structures in complex networks is facing a significant growth, as observations on real life graphs reveal that many social, biological, and technological networks are intrinsically divided into clusters.

¹The work of this author has been partially supported by INDAM-GNCS.

²The work of this author has been partially supported by the ERC Grant NOLEPRO.

Given a generic graph describing some kind of relationship among actors of a complex network, community detection problems basically consist in discovering and revealing the groups (if any) in which the network is subdivided.

Modularity matrices, the main subject of investigation of the present work, are a relevant tool in the development of a sound theoretical background of community detection. Despite a number of modularity matrices has been proposed so far, see e.g., [9] and the references therein, the original and most popular one was introduced by Newman and Girvan in [19] and is defined as a particular rank-one correction of the adjacency matrix. We shall refer to such matrix as the *Newman–Girvan* (or *unnormalized*) modularity matrix, and we will introduce consequently a normalized version of that matrix.

Spectral algorithms are widely applied to data clustering problems, including finding communities or partitions in graphs and networks. In the latter case, sign patterns in the entries of certain eigenvectors of Laplacian matrices are exploited to build vertex subsets, called *nodal domains*, which often yield excellent solutions to certain combinatorial problems related to the optimal partitioning of a given graph or network.

Analogously, nodal domains of modularity matrices play a crucial role in the community detection framework. A nodal domain theorem has been proved for these matrices [8, 9] showing the connectedness properties of nodal domains associated with their eigenvectors. The main results of this paper show that, under certain conditions, the nodal domains induced by eigenvectors corresponding to positive eigenvalues of the *normalized* modularity matrix have indeed positive modularity, that is they can be recognized as modules inside the graph. Moreover we prove two Cheeger-type inequalities for the cut-modularity providing a theoretical support to the common understanding that highly positive eigenvalues of modularity matrices are related with the possibility of subdividing the graph into communities.

The paper is organized as follows. After fixing hereafter our notation and preliminary results, in Section 2 we introduce with more detail the modularity based community detection problem, motivating our subsequent investigations. In Section 3 we discuss the unnormalized and normalized versions of the Newman–Girvan modularity matrix, summarizing some of their main structural properties. Subsequently, and we present our main results, concerning the relation between positive eigenvalues of the normalized modularity matrix and modules inside the graph. In particular in Section 4 we prove two Cheeger-type inequalities for the cut-modularity of the graph. Section 5 contains complementary results on modularity properties of nodal

domains corresponding to positive eigenvalues of the normalized modularity matrix. We devote a brief final section to few relevant concluding remarks.

1.1. Notations and preliminaries

Hereafter, we give a brief review of standard concepts and symbols from algebraic graph theory that we will use throughout the paper. We assume that $G = (V, E)$ is a simple connected graph, i.e., a finite, undirected, unweighted graph without multiple edges, where V and E are the vertex and edge sets. We always identify V with $\{1, \dots, n\}$. We denote adjacency of vertices x and y interchangeably as $x \sim y$ or $xy \in E$. Further definitions are listed hereafter:

- For any $i \in V$, let d_i denote its degree. Moreover, we let $d = (d_1, \dots, d_n)^T$, $\delta = (\sqrt{d_1}, \dots, \sqrt{d_n})^T$, $D = \text{Diag}(d_1, \dots, d_n)$.
- The symbols A and \mathcal{A} denote the adjacency matrix of G and its normalized counterpart, that is, $A \equiv (a_{ij})$ where $a_{ij} = 1$ if $ij \in E$, and $a_{ij} = 0$ otherwise; and $\mathcal{A} = D^{-1/2}AD^{-1/2}$. In particular, both A and \mathcal{A} are symmetric, irreducible, componentwise nonnegative matrices.
- $\mathbb{1}$ denotes the vector of all ones whose dimension depends on the context.
- The cardinality of a set S is denoted by $|S|$. In particular, $|V| = n$.
- For any $S \subseteq \{1, \dots, n\}$ let $\mathbb{1}_S$ be its *characteristic vector*, defined as $(\mathbb{1}_S)_i = 1$ if $i \in S$ and $(\mathbb{1}_S)_i = 0$ otherwise. Moreover, we denote by \bar{S} the complement $V \setminus S$, and let $\text{vol } S = \sum_{i \in S} d_i$ be the *volume* of S . Correspondingly, $\text{vol } V = \sum_{i \in V} d_i$ denotes the volume of the whole graph.
- For any subsets $S, T \subseteq V$ let

$$e(S, T) = \mathbb{1}_S^T A \mathbb{1}_T.$$

For simplicity, we use the shorthands $e_{\text{in}}(S) = e(S, S)$ and $e_{\text{out}}(S) = e(S, \bar{S})$, so that $e_{\text{in}}(S)$ is (twice) the number of inner-edges in S and $e_{\text{out}}(S)$ is the size of the edge-boundary of S . We have also

$$\text{vol } S = e_{\text{in}}(S) + e_{\text{out}}(S).$$

- A complete multipartite graph is a graph whose vertices can be partitioned into pairwise disjoint subsets V_1, \dots, V_k such that an edge exists if and only if the two extremes belong to different subsets, see e.g., [16]. In particular, if $k = n$ then G is a complete graph, while if $k = 2$ and V_1 is a singleton then G is a star.

2. The community detection problem

The discovery and description of communities in a graph is a central problem in modern graph analysis. Intuition suggests that a community (or cluster) in G should be a possibly connected group of nodes whose internal edges outnumber those with the rest of the network. However there is no formal definition of community. A survey of several proposed definitions of community can be found in [12], nonetheless as the author of that paper therein underlines, the global definition based on the modularity quality function is by far the most popular one. The modularity function was proposed by Newman and Girvan in [19] as a possible measure to quantify how much a given subset $S \subset V$ is a “good cluster”. They postulate that S is a cluster of nodes in G if the induced subgraph $G(S)$ contains more edges than expected, if edges were placed randomly. Thus, they introduce the modularity function $Q(S)$ to measure the difference between the actual and the expected number of edges in $G(S)$ so that a subset is a cluster if it has positive modularity. The precise definition is given by the following equivalent formulas:

$$Q(S) = e_{\text{in}}(S) - \frac{(\text{vol } S)^2}{\text{vol } V} = \frac{\text{vol } S \text{ vol } \bar{S}}{\text{vol } V} - e_{\text{out}}(S). \quad (1)$$

Note the equalities $Q(S) = Q(\bar{S})$ and $Q(V) = 0$. Undoubtedly, the modularity of a vertex set is one of the most efficient indicators of its consistency as a community in G . For that reason, it is common practice to adopt the following definition:

Definition 2.1. A subgraph of G is a *module* if its vertex set S has positive modularity. If no ambiguity may occur, S is called a module itself.

The usefulness of the previous definition lies in the fact that, in practice, if $G(S)$ is a connected module whose size is significant, then it can be recognized as a community.

Definition 2.1 leads naturally to an efficient measure of a partitioning of G into modules. Indeed, let S_1, \dots, S_k be a partition of V into pairwise

disjoint subsets. The (*normalized*) *modularity* of S_1, \dots, S_k is defined as

$$q(S_1, \dots, S_k) = \frac{1}{\text{vol } V} \sum_{i=1}^k Q(S_i). \quad (2)$$

The normalization factor $1/\text{vol } V$ is somehow conventional. It has been introduced in [17, 19] to settle the value of q in a range independent on G and k and for compatibility with previous works.

The problem of partitioning a graph into an arbitrary number of subgraphs whose overall modularity is maximized has received a considerable attention, not only in its applicative and computational aspects but also from the graph-theoretic point of view [6, 14]. The main contributions we propose in this work shall deal with the *cut version* of the community detection problem, that is the problem of finding a subset $S \subseteq V$ having maximal modularity (uniqueness is not ensured in the general case). To this end, it is worth to define the *cut-modularity of the graph G* as the quantity

$$q_G^{Cut} = \max_{S \subseteq V} q(S, \bar{S}) = \frac{2}{\text{vol } V} \max_{S \subseteq V} Q(S). \quad (3)$$

It is well known that the optimization of the modularity function (2) presents some drawbacks when employed for finding a partitioning of G into modules, since small clusters tend to be subsumed by larger ones. Among the many techniques and variants of the Newman–Girvan modularity that have been devised to tackle this issue, which is widely known as *resolution limit*, here we borrow from [1] two weighted versions of the modularity function that play a relevant role in the subsequent discussion:

- The *relative modularity* of $S \subseteq V$ is $Q_{\text{rel}}(S) = Q(S)/|S|$. This definition is naturally extended to the cut $\{S, \bar{S}\}$ as

$$q_{\text{rel}}(S, \bar{S}) = Q_{\text{rel}}(S) + Q_{\text{rel}}(\bar{S}) = Q(S) \frac{n}{|S||\bar{S}|}, \quad (4)$$

which, in turn, leads to the definition of the *relative cut-modularity of G*

$$q_G^{RCut} = \max_{S \subseteq V} q_{\text{rel}}(S, \bar{S}).$$

- The *normalized modularity* of $S \subseteq V$ is defined as $Q_{\text{norm}}(S) = Q(S)/\text{vol } S$ and that definition can be extended to the cut $\{S, \bar{S}\}$ as

$$q_{\text{norm}}(S, \bar{S}) = Q_{\text{norm}}(S) + Q_{\text{norm}}(\bar{S}) = Q(S) \frac{\text{vol } V}{\text{vol } S \text{vol } \bar{S}}. \quad (5)$$

As before we define the *normalized cut-modularity* of the graph G as

$$q_G^{NCut} = \max_{S \subseteq V} q_{\text{norm}}(S, \bar{S}).$$

Straightforward computations ensure

$$\frac{2 q_G^{RCut}}{n d_{\max}} \leq q_G^{Cut} \leq \frac{q_G^{RCut}}{2}, \quad \frac{2 q_G^{NCut}}{\text{vol } V} \leq q_G^{Cut} \leq \frac{q_G^{NCut}}{2}.$$

3. Modularity matrices and their properties

The probably best known methods for detecting a subset whose modularity well approximates the cut-modularity of G are based on the idea of spectral partitioning and are related with an important rank-one correction of the adjacency matrix, known as *the Newman–Girvan modularity matrix*. In analogy with the graph Laplacians, in this section we define two different modularity matrices, describing a number of relevant structural properties.

3.1. The Newman–Girvan modularity matrix

Given a graph G and the associated adjacency matrix A , let $d = A\mathbf{1}$ be the degree vector of G , and $\text{vol } V = \sum_i d_i$ be its volume. The *unnormalized modularity matrix* of G has been introduced in [17] as the following rank one perturbation of A :

$$M = A - \frac{1}{\text{vol } V} dd^T. \quad (6)$$

For any $S \subseteq V$ let $\mathbf{1}_S$ be its characteristic vector: $(\mathbf{1}_S)_i = 1$ if $i \in S$ and $(\mathbf{1}_S)_i = 0$ otherwise. With the help of these notations we can express $Q(S)$ as

$$Q(S) = \mathbf{1}_S^T M \mathbf{1}_S \quad (7)$$

The following proposition summarizes some basics properties of M :

Proposition 3.1. *The matrix M satisfies the following properties:*

1. M is symmetric and $\mathbf{1} \in \ker(M)$.
2. If $m_1 \geq \dots \geq m_n$ are the eigenvalues of M and $\alpha_1 \geq \dots \geq \alpha_n$ those of A , then $\alpha_1 \geq m_1 \geq \alpha_2 \geq m_2 \geq \dots \geq \alpha_n \geq m_n$.
3. 0 is a simple eigenvalue of M if and only if A is nonsingular.
4. The rightmost eigenvalue of M is nonnegative, and is zero if and only if G is a complete multipartite graph.

Proof. Point 1 is revealed by a direct computation. Point 2 is a direct consequence of the variational characterization of the eigenvalues of symmetric matrices, see e.g., [22]. To show point 3 we observe that the multiplicity of the zero eigenvalue of M is one plus the dimension of the kernel of A . Indeed consider the diagonal matrix $\Delta = \text{Diag}(1/\sqrt{d_1}, \dots, 1/\sqrt{d_n})$ and let $\delta = \Delta d$. Then $\Delta M \Delta \delta = 0$ and $\Delta A \Delta \delta = \delta$. Therefore the multiplicity of the zero eigenvalue of $\Delta M \Delta$ is the multiplicity of the zero eigenvalue of $\Delta A \Delta$ plus one. This proves point 3 as the multiplicity of 0 is invariant under matrix congruences. Point 4 is a rephrasing of Theorem 1.1 in [16]. See also [2, Thm. 11]. \square

The modularity matrix M is at the basis of many spectral methods for community detection, and the eigenstructure of M can be used to describe clustering properties of graphs. In particular, the nodal domains associated to its principal eigenvectors cover a special role, as they are often good candidates for leading modules inside G . A number of results relating algebraic properties of M to communities in G have appeared in recent literature [1, 2, 8, 9, 16], the forthcoming Theorem 3.2 summarizes those among them which to our opinion are most relevant.

As it often plays a special role in the algebraic analysis of the modular structure of G , the rightmost nonzero eigenvalue of M deserves a special symbol, borrowed from [8] and therein named *algebraic modularity*:

$$m_G = \max_{\substack{v \in \mathbb{R}^n \\ v^T \mathbf{1} = 0}} \frac{v^T M v}{v^T v}. \quad (8)$$

Already at this stage intuition suggests that a close relation should exist between m_G and the cut-modularity (3), and that the subsets $S \subseteq V$ having positive modularity should be related with positive eigenvalues of M . The following theorem summarizes some important eigenproperties of M that have been proven in recent literature, see in particular, [2, 8, 16].

Theorem 3.2. *The matrix M satisfies the following properties:*

1. $m_G < \rho(A)$ and, if d is not an eigenvector of A , then m_G is simple.
2. If G is not a complete graph or a complete multipartite graph then $m_G = \lambda_1(M)$, the rightmost eigenvalue of M , and is positive. If G is a star then $m_G = \lambda_2(M)$, the second rightmost eigenvalue of M , and is negative. Otherwise (that is, if G is a complete graph or a complete multipartite graph which is not a star) $m_G = 0$.
3. Let $\langle d \rangle = \text{vol } V/n$ be the average degree of G , then $m_G \geq 2 \langle d \rangle q_G^{\text{Cut}}$.

4. Let $\{S_1, \dots, S_k\}$ be a partition that maximizes the quantity in (2), which has minimal cardinality, and which is made up entirely by modules. Then $k - 1$ does not exceed the number of positive eigenvalues of M .
5. Let u be an eigenvector associated with m_G such that $d^T u \geq 0$. If m_G is simple and it is not an eigenvalue of A then the subgraph induced by the subset $S_+ = \{i \mid u_i \geq 0\}$ is connected.

For any $S \subseteq V$ let $v_S = \mathbb{1}_S - \frac{|S|}{n} \mathbb{1}$. The following identities are readily obtained:

$$v_S^T \mathbb{1} = 0, \quad v_S^T v_S = \frac{|S| |\bar{S}|}{n}, \quad v_S^T M v_S = Q(S), \quad q_{\text{rel}}(S, \bar{S}) = \frac{v_S^T M v_S}{v_S^T v_S}.$$

Hence, the combinatorial problem of finding the cut $\{S, \bar{S}\}$ with largest relative modularity has a natural continuous relaxation in the maximization of the Rayleigh quotient $v^T M v / v^T v$ over the subspace orthogonal to $\mathbb{1}$, that is, the algebraic modularity defined in (8). We have the immediate consequence

$$q_G^{\text{RCut}} \leq m_G.$$

3.2. The normalized modularity matrix

In analogy with the renowned normalized Laplacian matrix of a graph, we let $\mathcal{A} = D^{-1/2} A D^{-1/2}$ be the normalized adjacency matrix and define the *normalized modularity matrix* of G as

$$\mathcal{M} = D^{-1/2} M D^{-1/2} = \mathcal{A} - \frac{1}{\text{vol } V} \delta \delta^T$$

where $\delta = (\sqrt{d_1}, \dots, \sqrt{d_n})^T$ and M is as in (6). The matrix \mathcal{M} appeared recently in the community detection literature, and in various other network related questions as the analysis of quasi-randomness properties of graphs with given degree sequences, see [1, 4, 9] and [3, Chap. 5]. Several basics properties of \mathcal{M} can be immediately observed; we collect some of them hereafter.

Proposition 3.3. *The matrix \mathcal{M} satisfies the following properties:*

1. \mathcal{M} has a zero eigenvalue with corresponding eigenvector δ .
2. The matrices \mathcal{M} and \mathcal{A} coincide over the space orthogonal to δ . That is, $\mathcal{M}v = \mathcal{A}v$ for all $v \in \langle \delta \rangle^\perp$.
3. The eigenvalues of \mathcal{M} belong to the interval $[-1, 1]$. Moreover, 0 is a simple eigenvalue of \mathcal{M} if and only if \mathcal{A} is nonsingular.

4. If G is connected then 1 is not an eigenvalue of \mathcal{M} . Furthermore, if G is not bipartite then -1 is not an eigenvalue of \mathcal{M} .

Proof. Straightforward computations show that $\mathcal{A}\delta = \delta$ and $\mathcal{M}\delta = 0$. Since $\mathcal{A} \geq O$ and $\delta \geq 0$, Perron–Frobenius theory leads us to deduce that $\rho(\mathcal{A}) = 1$ is an eigenvalue of \mathcal{A} . Therefore, if $\mathcal{A} = \sum_{i=1}^n \lambda_i q_i q_i^T$ is a spectral decomposition of \mathcal{A} with the eigenvalues in nonincreasing order, $\lambda_1 \geq \dots \geq \lambda_n$, then we can assume $\lambda_1 = 1$, $|\lambda_i| \leq 1$ for $i > 1$, and q_1 parallel to δ . In particular, $\delta\delta^T/\text{vol } V$ is the orthogonal projector on the eigenspace spanned by q_1 , since $\delta^T\delta = \text{vol } V$. Consequently, $\mathcal{M} = \sum_{i=2}^n \lambda_i q_i q_i^T$ is a spectral decomposition of \mathcal{M} and we easily deduce points 2 and 3. Incidentally, this proves that \mathcal{M} and \mathcal{A} are simultaneously diagonalizable. If G is connected then \mathcal{A} is irreducible and λ_1 is simple, that is $1 > \lambda_2$. Furthermore, if G is not bipartite then \mathcal{A} is also primitive and $|\lambda_i| < 1$ for $i > 1$, and the proof is complete. \square

The normalized modularity (5) of a cut $\{S, \bar{S}\}$ can be naturally defined in terms of \mathcal{M} . In fact, given any $S \subseteq V$, consider the vector

$$v_S = D^{1/2}(\mathbb{1}_S - c\mathbb{1}), \quad c = \text{vol } S / \text{vol } V. \quad (9)$$

Simple computations prove that

$$\delta^T v_S = 0, \quad v_S^T v_S = \frac{\text{vol } S \text{ vol } \bar{S}}{\text{vol } V}.$$

Moreover,

$$\frac{v_S^T \mathcal{M} v_S}{v_S^T v_S} = \frac{(\mathbb{1}_S - c\mathbb{1})^T \mathcal{M} (\mathbb{1}_S - c\mathbb{1})}{v_S^T v_S} = \frac{\mathbb{1}_S^T \mathcal{M} \mathbb{1}_S}{\text{vol } S \text{ vol } \bar{S}} \text{vol } V = q_{\text{norm}}(S, \bar{S}).$$

It follows that the problem of computing the normalized cut-modularity of G can be stated in terms of \mathcal{M} . Indeed, if \mathcal{V}_n is the set of n -vectors having the form (9) for some $S \subset V$, then v_S is a generic vector in \mathcal{V}_n , implying that

$$q_G^{NCut} = \max_{v \in \mathcal{V}_n} \frac{v^T \mathcal{M} v}{v^T v} \quad (10)$$

and of course, if \hat{v} is the vector realizing the maximum in (10), then the set $\hat{S} = \{i \mid \hat{v}_i > 0\}$ defines the optimal cut. As for the unnormalized case, it is worth defining the *normalized algebraic modularity*:

$$\mu_G = \max_{\substack{v \in \mathbb{R}^n \\ v^T \delta = 0}} \frac{v^T \mathcal{M} v}{v^T v}. \quad (11)$$

Note that (11) is a relaxed version of (10). In particular,

$$q_G^{NCut} \leq \mu_G. \quad (12)$$

Since \mathcal{M} is real symmetric we immediately note that μ_G coincides with the largest eigenvalue of \mathcal{M} after deflation of the invariant subspace spanned by δ . Therefore, if $-1 \leq \mu_n \leq \dots \leq \mu_1 \leq 1$ are the eigenvalues of \mathcal{M} , then $\mu_1 = \max\{0, \mu_G\}$. Furthermore, since M and \mathcal{M} are related by a congruence transform, point 2 of Theorem 3.2 leads us to the following result:

Corollary 3.4. *If G is not a star then $\mu_G = \mu_1$, the rightmost eigenvalue of \mathcal{M} . Moreover, $\mu_G > 0$ if and only if G is not a complete graph or a complete multipartite graph.*

4. Cheeger-type inequalities

As we already discussed above, both heuristics and intuition suggest that μ_G quantifies the cut-modularity of the graph, and can be used to approximate q_G^{NCut} . While the upper bound $q_G^{NCut} \leq \mu_G$ has been shown in (12) by simple arguments, a converse relation, bounding q_G^{NCut} from below in terms of μ_G , is not that easy. In fact, there it is possible that $\mu_G > 0$ while $q_G^{NCut} < 0$, as shown experimentally in [2]. Theorems 4.1 and 4.3 contribute to this question stating lower (and upper) bounds of q_G^{NCut} in terms of spectral properties of \mathcal{M} .

The conductance (or sparsity, or Cheeger constant) h_G is one of the best known topological invariants of a graph G . For $S \subset V$ let

$$h(S) = \frac{e_{\text{out}}(S)}{\min\{\text{vol } S, \text{vol } \bar{S}\}},$$

the so-called conductance of S . Then, the conductance of G is defined as $h_G = \min_{S \subset V} h(S)$. Such quantity plays a fundamental role in graph partitioning problems [18, Chap. 11], in isoperimetric problems [3, Chap. 2], mixing properties of random walks, combinatorics, and in various other areas of mathematics and computer science. A renowned result in graph theory, known as *Cheeger inequality*, relates the conductance of G and the smallest positive eigenvalue of the normalized Laplacian matrix $\mathcal{L} = I - \mathcal{A}$.

If $0 = \lambda_1 < \lambda_2 \leq \dots \leq \lambda_n \leq 2$ are the eigenvalues of \mathcal{L} , the Cheeger inequality states that

$$\frac{1}{2}\lambda_2 \leq h_G \leq \sqrt{2\lambda_2}.$$

Actually, Chung [3] improved the upper bound to $h_G \leq \sqrt{\lambda_2(2 - \lambda_2)}$. Let v be an eigenvector of \mathcal{L} corresponding to λ_2 and consider the equality

$\mathcal{L} = I - \mathcal{A} = I - \mathcal{M} + \delta\delta^T/\delta^T\delta$. Since $\mathcal{L}\delta = 0$, we have $\delta^Tv = 0$. By Courant's minimax principle and (11),

$$\lambda_2 = \min_{v:\delta^Tv=0} \frac{v^T\mathcal{L}v}{v^Tv} = 1 - \max_{v:\delta^Tv=0} \frac{v^T\mathcal{M}v}{v^Tv} = 1 - \mu_G.$$

In particular, from Corollary 3.4 we obtain that, if G is not a star then $1 - \lambda_2$ is the rightmost eigenvalue of M . A direct application of the Cheeger inequality yields the following estimates for q_G^{NCut} .

Theorem 4.1. *Let μ_1 be the rightmost eigenvalue of \mathcal{M} . If G is not a star then*

$$1 - 2\sqrt{1 - \mu_1^2} \leq q_G^{NCut} \leq \mu_1.$$

Proof. Recalling (1) and (5), we have

$$\begin{aligned} q_{\text{norm}}(S, \bar{S}) &= \frac{\text{vol } V}{\text{vol } S \text{vol } \bar{S}} Q(S) \\ &= 1 - \frac{\text{vol } V}{\text{vol } S \text{vol } \bar{S}} e_{\text{out}}(S) \geq 1 - 2h(S), \end{aligned}$$

since $\text{vol } V / \text{vol } S \text{vol } \bar{S} \leq 2 / \min\{\text{vol } S, \text{vol } \bar{S}\}$. By maximizing over S we eventually get

$$q_G^{NCut} = \max_{S \subset V} q_{\text{norm}}(S, \bar{S}) \geq 1 - 2h_G \geq 1 - 2\sqrt{(1 - \mu_G)(1 + \mu_G)}.$$

By hypothesis, $\mu_G = \mu_1$. The upper bound comes from (12). \square

Extensive research on Cheeger-type results by many authors suggests that no substantial improvements on the lower bound in Theorem 4.1 can be obtained without additional information on G , although explicit examples of graph sequences proving optimality of that bound are not known. However, the forthcoming result shows that, almost surely, $1 - \mu_1$ can be a much better estimate to $1 - q_G^{NCut}$ than expected, in particular, when the entries of an eigenvector of μ_1 cluster around two values. We will make use of the following lemma, whose simple proof is omitted for brevity:

Lemma 4.2. *If $\sum_{i=1}^n \alpha_i = 0$ then $\sum_{i:\alpha_i>0} \alpha_i = \frac{1}{2} \sum_{i=1}^n |\alpha_i|$.*

Theorem 4.3. *Let μ_1 be the rightmost eigenvalue of \mathcal{M} . Suppose that μ_1 has an eigenvector x without zero entries. Then there exists a constant $C > 0$, not depending on μ_1 , such that*

$$1 - C(1 - \mu_1) \leq q_G^{NCut}.$$

Proof. Let v be an eigenvector of \mathcal{M} corresponding to μ_1 and let $z = D^{-1/2}v$. Note that v is orthogonal to the vector $\delta = (\sqrt{d_1}, \dots, \sqrt{d_n})^T$, since the latter is an eigenvector of \mathcal{M} associated to 0. Consequently, z is orthogonal to the degree vector: $d^T z = \delta^T D^{1/2} z = \delta^T v = 0$. Hence,

$$\mu = \frac{v^T \mathcal{M} v}{v^T v} = \frac{v^T \mathcal{A} v}{v^T v} = \frac{z^T A v}{z^T D z} = 1 - \frac{z^T L z}{z^T D z},$$

where $L = D - A$ is the Laplacian matrix of G . We have

$$z^T L z = \sum_{ij \in E} (z_i - z_j)^2,$$

where the sum runs over the edges of the graph, each edge being counted only once. On the other hand,

$$z^T D z = \sum_{i=1}^n d_i z_i^2.$$

For notational simplicity, we use the shorthands $s = \text{vol } S$, $\bar{s} = \text{vol } \bar{S}$, and $\nu = s + \bar{s} = \text{vol } V$. Consider the nodal domain $S = \{i : v_i \geq 0\}$ and let x be the step vector $x = p\mathbb{1}_S + q\mathbb{1}_{\bar{S}}$ which minimizes the weighted distance

$$\|D^{1/2}(x - z)\|_2^2 = \sum_{i=1}^n d_i (x_i - z_i)^2 = \sum_{i \in S} d_i (p - z_i)^2 + \sum_{i \in \bar{S}} d_i (q - z_i)^2.$$

Simple computations show that the minimum is attained when

$$p = \left(\sum_{i \in S} d_i z_i \right) / s, \quad q = \left(\sum_{i \in \bar{S}} d_i z_i \right) / \bar{s}.$$

Observe that p and q are weighted averages of the values z_i for $i \in S$ and $i \in \bar{S}$, respectively. With the notation $c = \sum_{i \in S} d_i z_i$, from the orthogonality condition $d^T z = 0$ and Lemma 4.2 we deduce the simpler formulas $p = c/s$ and $q = -c/\bar{s}$. For later reference, we remark the identities

$$p - q = \frac{c\nu}{s\bar{s}}, \quad p^2 s + q^2 \bar{s} = \nu \frac{(c\nu)^2}{(s\bar{s})^2}. \quad (13)$$

Incidentally, we note that, apart of a constant, the vector $D^{1/2}x$ coincides with the vector in (9). Moreover, it is not hard to recognize that, if G is disconnected then the vector $D^{1/2}x$ is an eigenvector of \mathcal{M} associated to the

eigenvalue 1. Our subsequent arguments are based on the intuition that, if z is a small perturbation of x then S is weakly linked to \bar{S} . Let $r \geq 1$ be a number such that

$$r^{-1} \leq z_i/x_i \leq r, \quad i = 1, \dots, n.$$

In fact, if $z_i > 0$ then $x_i = p > 0$, whereas $z_i < 0$ implies $x_i = q < 0$. Hence, if $ij \in E$ is an edge joining a node in S with a node in \bar{S} we have $|z_i - z_j| \geq (p - q)/r$. Consequently,

$$z^T L z = \sum_{ij \in E} (z_i - z_j)^2 \geq r^{-2} (p - q)^2 e_{\text{out}}(S),$$

by neglecting all contributions from edges lying entirely inside S or \bar{S} . Moreover,

$$z^T D z = \sum_{i=1}^n d_i z_i^2 \leq r^2 \left(\sum_{i \in S} p^2 d_i + \sum_{i \in \bar{S}} q^2 d_i \right) = r^2 (p^2 s + q^2 \bar{s}).$$

Consider the equality $e_{\text{out}}(S) = (1 - q_{\text{norm}}(S, \bar{S})) s \bar{s} / \nu$. Using (13) and simplifying we get

$$1 - \mu = \frac{z^T L z}{z^T D z} \geq \frac{1}{r^4 \nu} e_{\text{out}}(S) = \frac{s \bar{s}}{r^4 \nu^2} (1 - q_{\text{norm}}(S, \bar{S})) \geq \frac{1}{4 r^4} (1 - q_G^{NCut}),$$

owing to $s \bar{s} / \nu^2 \geq \frac{1}{4}$. \square

5. Modules from nodal domains

Theorems 4.1 and 4.3 state in particular that if μ_G is sufficiently close to 1, then the cut-modularity of G is positive and thus there exists a bipartition of V into $\{S, \bar{S}\}$ such that both $G(S)$ and $G(\bar{S})$ are modules. Of course such bipartition is not unique in the general case. The forthcoming theorems strengthen this claim by showing that, if a positive eigenvalue μ of \mathcal{M} is large enough, then we can explicitly exhibit a cut $\{S, \bar{S}\}$ with positive modularity, by defining it in terms of a nodal domain induced by an eigenvector corresponding to μ .

Given a nonzero vector $v \in \mathbb{R}^n$ the subgraph $G(S)$ induced by the set $S = \{i : v_i \geq 0\}$ is a *nodal domain* of v [5, 7]. This fundamental definition admits obvious variations (for example, inequality can be strict, or reversed) and, since the seminal papers by Fiedler [10, 11], it has become a major

tool for spectral methods in community detection and graph partitioning [17, 20, 21]. Indeed, nodal domains of eigenvectors of modularity matrices are commonly utilized in order to localize modules inside a network. If v is an eigenvector corresponding to μ_G , it has been shown in [9] that $S = \{i : v_i \geq 0\}$ induces a connected subgraph $G(S)$. The following Theorems 5.1 and 5.2 provide additional information on $G(S)$ as they show that, if μ_G is large enough, then the subgraph $G(S)$ is a module.

Theorem 5.1. *Let v be a normalized eigenvector of \mathcal{M} corresponding to a positive eigenvalue μ , that is, $\mathcal{M}v = \mu v$ with $\|v\|_2 = 1$. Let $S = \{i \mid v_i \geq 0\}$. If*

$$\mu > \frac{(\text{vol } S)^2 + (\text{vol } \bar{S})^2}{\text{vol } V} \max_{i \in V} \frac{v_i^2}{d_i}$$

then $Q(S) > 0$.

Proof. Recalling Proposition 3.3, we have that v is orthogonal to δ , which implies in turn $\mathcal{M}v = \mathcal{A}v$ and $\mu = v^T \mathcal{M}v = v^T \mathcal{A}v$. Define the set $\mathcal{I}_+ = (S \times S) \cup (\bar{S} \times \bar{S})$. Note that $v_i v_j \geq 0$ whenever $(i, j) \in \mathcal{I}_+$. Using entrywise nonnegativity of \mathcal{A} we obtain

$$\mu = v^T \mathcal{A}v \leq \sum_{(i,j) \in \mathcal{I}_+} v_i v_j \mathcal{A}_{ij} \leq \left(\max_{i \in V} \frac{|v_i|}{\delta_i} \right)^2 \sum_{(i,j) \in \mathcal{I}_+} \delta_i \delta_j \mathcal{A}_{ij}.$$

Since $\delta_i \delta_j \mathcal{A}_{ij} = A_{ij}$, the rightmost summations yield

$$\sum_{(i,j) \in \mathcal{I}_+} A_{ij} = \mathbb{1}_{\bar{S}}^T A \mathbb{1}_S + \mathbb{1}_{\bar{S}}^T A \mathbb{1}_{\bar{S}} = e_{\text{in}}(S) + e_{\text{in}}(\bar{S}).$$

Let us set $C^2 = (\max_{i \in V} |v_i|/\delta_i)^2$. Owing to the equalities $Q(S) = e_{\text{in}}(S) - (\text{vol } S)^2/\text{vol } V$ and $Q(S) = Q(\bar{S})$ we have

$$\mu \leq C^2 (e_{\text{in}}(S) + e_{\text{in}}(\bar{S})) = C^2 \left(2Q(S) + \frac{(\text{vol } S)^2 + (\text{vol } \bar{S})^2}{\text{vol } V} \right).$$

By rearranging terms,

$$2C^2 Q(S) \geq \mu - C^2 \frac{(\text{vol } S)^2 + (\text{vol } \bar{S})^2}{\text{vol } V},$$

and the claim follows. \square

With respect to the quantity $\max_i v_i^2/d_i$ appearing in the preceding theorem, consider that if G is k -regular (that is, $d_i = k$ for every $i \in V$) then $v_i = n^{-\frac{1}{2}}$ and $\text{vol } V = kn$. After simple passages the aforementioned lower bound for μ becomes $(|S|^2 + |\bar{S}|^2)/n^2$, a number which is strictly smaller than 1.

Theorem 5.2. *Let v be any real eigenvector of \mathcal{M} corresponding to a positive eigenvalue μ , that is, $\mathcal{M}v = \mu v$. Let $S = \{i \mid v_i \geq 0\}$ and let $\cos \theta$ be the cosine of the acute angle between the vectors $|v| = (|v_1|, \dots, |v_n|)^T$ and $\delta = (\sqrt{d_1}, \dots, \sqrt{d_n})^T$. If*

$$\mu + 1 > 4 \frac{\text{vol } S \text{ vol } \bar{S}}{(\text{vol } V)^2} \frac{1}{\cos^2 \theta}$$

then $Q(S) > 0$.

Proof. Let $s = D^{1/2} \mathbb{1}_S$, that is

$$s_i = \begin{cases} \delta_i & v_i \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Observe that $\|s\|_2^2 = \sum_{i \in S} d_i = \text{vol } S$ and $\delta^T s = \text{vol } S$ too. Since v is orthogonal to $\delta = (\sqrt{d_1}, \dots, \sqrt{d_n})^T$, there exist scalars α, β, γ such that we have the orthogonal decomposition

$$s = \alpha \frac{1}{\|\delta\|_2} \delta + \beta \frac{1}{\|v\|_2} v + \gamma w \quad (14)$$

for some normalized vector $w \in \mathbb{R}^n$ orthogonal to both δ and v . The coefficients in (14) own the following explicit formulas:

$$\alpha = \frac{1}{\|\delta\|_2} \delta^T s = \frac{\text{vol } S}{\sqrt{\text{vol } V}}, \quad \beta = \frac{v^T s}{\|v\|_2},$$

and moreover,

$$\begin{aligned} \gamma^2 &= \|s\|_2^2 - \alpha^2 - \beta^2 = \text{vol } S - \frac{(\text{vol } S)^2}{\text{vol } V} - \beta^2 \\ &= \frac{\text{vol } S \text{ vol } \bar{S}}{\text{vol } V} - \beta^2. \end{aligned}$$

Owing to the fact that the spectrum of \mathcal{M} is included in $[-1, 1]$ and the assumption $\|w\|_2 = 1$ we have $w^T \mathcal{M} w \geq -1$. Hence, from (14) we obtain

$$\begin{aligned} Q(S) &= \mathbb{1}_S^T \mathcal{M} \mathbb{1}_S = s^T \mathcal{M} s \\ &\geq \alpha^2 \cdot 0 + \beta^2 \mu - \gamma^2 = \beta^2 (\mu + 1) - \frac{\text{vol } S \text{ vol } \bar{S}}{\text{vol } V}. \end{aligned}$$

Thus, if

$$\mu + 1 > \frac{\text{vol } S \text{ vol } \bar{S}}{\beta^2 \text{ vol } V}$$

then $Q(S) > 0$. Moreover, using the orthogonality $\delta^T v = 0$ and Lemma 4.2 we obtain

$$\cos \theta = \frac{\sum_{i \in V} \delta_i |v_i|}{\|v\|_2 \|\delta\|_2} = \frac{2 \sum_{i \in S} \delta_i v_i}{\|v\|_2 \sqrt{\text{vol } V}} = 2 \frac{v^T s}{\|v\|_2 \sqrt{\text{vol } V}},$$

whence $\beta = \frac{1}{2}(\cos \theta) \sqrt{\text{vol } V}$ and the proof is complete. \square

From the straightforward bound

$$\text{vol } S \text{ vol } \bar{S} / (\text{vol } V)^2 \leq \frac{1}{4}$$

and the equality $\cos^{-2} \theta - 1 = \tan^2 \theta$, we derive the following condition.

Corollary 5.3. *In the same notations of Theorem 5.2, if $\mu > \tan^2 \theta$ then $Q(S) > 0$.*

6. Concluding remarks

Community detection is a major task in modern complex network analysis and the matrix approach to such problem is quite popular and powerful. In this work we formulate the modularity of a cut in terms of a quadratic form associated with the normalized modularity matrix, and we provide theoretical supports to the common understanding that highly positive eigenvalues of the normalized modularity matrix imply the presence of communities in G . In particular we show that, if that matrix has an eigenvalue close to 1 then the nodal domains corresponding to that eigenvalue have positive modularity and, moreover, can produce good estimates of the optimal cut-modularity.

As recent advances in spectral graph theory have shown higher order Cheeger inequalities in terms of higher order eigenvalues of the graph Laplacian [13, 15], we believe that deeper spectral based investigations could reveal more precise relations between the magnitude and the number of positive eigenvalues of the modularity matrices and the presence of communities in the network.

References

- [1] M. Bolla. Penalized versions of the Newman–Girvan modularity and their relation to normalized cuts and k -means clustering. *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.*, 84:1–12, 2011.
- [2] Marianna Bolla, Brian Bullins, Sorathan Chaturapruek, Shiwen Chen, and Katalin Friedl. Spectral properties of modularity matrices. *Linear Algebra Appl.*, 473:359–376, 2015.
- [3] F. R. K. Chung. *Spectral Graph Theory*, volume 92 of *CBMS Regional Conference Series in Mathematics*. AMS, 1997.
- [4] Fan Chung and Ron Graham. Quasi-random graphs with given degree sequences. *Random Structures Algorithms*, 32(1):1–19, 2008.
- [5] E. B. Davies, G. M. L. Gladwell, J. Leydold, and P. F. Stadler. Discrete nodal domain theorems. *Linear Algebra Appl.*, 336:51–60, 2001.
- [6] Fabien de Montgolfier, Mauricio Soto, and Laurent Viennot. Asymptotic modularity of some graph classes. In *Algorithms and computation*, volume 7074 of *Lecture Notes in Comput. Sci.*, pages 435–444. Springer, Heidelberg, 2011.
- [7] A. M. Duval and V. Reiner. Perron–Frobenius type results and discrete versions of nodal domain theorems. *Linear Algebra Appl.*, 294:259–268, 1999.
- [8] D. Fasino and F. Tudisco. An algebraic analysis of the graph modularity. *SIAM J. Matrix Anal. Appl.*, 35(3):997–1018, 2014.
- [9] D. Fasino and F. Tudisco. Generalized modularity matrices. *Linear Algebra Appl.*, (to appear), 2015.
- [10] M. Fiedler. Algebraic connectivity of graphs. *Czechoslovak Mathematical Journal*, 23:298–305, 1973.
- [11] M. Fiedler. A property of eigenvectors of nonnegative symmetric matrices and its application to graph theory. *Czechoslovak Mathematical Journal*, 25(100):619–633, 1974.
- [12] S. Fortunato. Community detection in graphs. *Physics Reports*, 486:75–174, 2010.

- [13] M. Hein and F. Tudisco. Multi-way Cheeger inequalities for the graph p -Laplacian. *preprint*, 2015.
- [14] Ath. Kehagias and L. Pitsoulis. Bad communities with high modularity. *Eur. Phys. J. B*, 86(7):Art. 330, 11, 2013.
- [15] J. R. Lee, S. O. Gharan, and L. Trevisan. Multi-way spectral partitioning and higher-order Cheeger inequalities. In *Proceedings of the Forty-fourth Annual ACM Symposium on Theory of Computing*, STOC '12, pages 1117–1130, New York, NY, USA, 2012. ACM.
- [16] S. Majstorovic and D. Stevanovic. A note on graphs whose largest eigenvalues of the modularity matrix equals zero. *Electronic Journal of Linear Algebra*, 27:611–618, 2014.
- [17] M. E. J. Newman. Finding community structure in networks using the eigenvectors of matrices. *Phys. Rev. E*, 69:321–330, 2006.
- [18] M. E. J. Newman. *Networks: An Introduction*. OUP Oxford, 2010.
- [19] M. E. J. Newman and M. Girvan. Finding and evaluating community structure in networks. *Phys. Rev. E*, 69(026113), 2004.
- [20] D. L. Powers. Graph partitioning by eigenvectors. *Linear Algebra Appl.*, 101:121–133, 1988.
- [21] S. E. Schaeffer. Graph clustering. *Computer Science Review*, 1(1):27 – 64, 2007.
- [22] J. H. Wilkinson. *The algebraic eigenvalue problem*. Clarendon Press, Oxford University Press, Walton Street, 1965.